

- В процессе жизнедеятельности у агентов расходуется энергия. Если запас энергии у агента исчерпывается, он умирает и превращается в пищу определенного типа для других агентов.

В ходе симуляции при варьировании начальных данных наблюдались следующее. При исчерпаемом первичном источнике пищи популяция лавинно размножалась пока не достигала кризиса. Затем преимущество получали особи, получившие в ходе мутации способность есть «тела» погибших агентов. При неисчерпаемом первичном источнике пищи наблюдалось свободное сосуществование нескольких популяций, отличающихся типом пищи, употребляемым в качестве источника энергии.

1. R. E. Lenski et al, *Nature*, 461, 1243 (2009).
2. R. E. Lenski et al, *Nature*, 551, 45 (2017).

ВЫБОР МЕТОДА ДЛЯ НАКОПЛЕНИЯ ТЕРМИНОВ ИЗ ИНТЕРЕСУЮЩЕЙ СУБЪЕКТА ПРЕДМЕТНОЙ ОБЛАСТИ

Ходенева М.А.^{1*}, Кудрявцев А.Г.¹

¹⁾ Уральский федеральный университет имени первого Президента России Б.Н. Ельцина, г. Екатеринбург, Россия

*E-mail: masha.hodeneva@mail.ru

CHOICE OF THE METHOD FOR ACCUMULATION OF TERMS FROM THE SUBJECT AREA OF INTERESTS THE USER

Hodeneva M.A.^{1*}, Kudryavtsev A.G.¹

¹⁾ Ural Federal University, Yekaterinburg, Russia

Annotation. Review was undertaken of systems for extracting a terminological dictionary favouring the interests of the user, pointed out certain deficiencies in the systems and proposed solutions to overcome those obstacles.

В процессе разработки фильтра текстов по семантической схожести с интересами субъекта были определены его основные структурные составляющие, в число которых входит система подготовки информации по терминам. Данная система предназначена для обеспечения тестирования субъекта с целью извлечения информации о его интересах.

Был проведен литературно-аналитический обзор систем подготовки информации по терминам, в результате которого были выявлены два аналога:

1. Системы, основанные на использовании словаря терминов предметной области [1].

2. Системы, основанные на провоцировании выдачи текстов субъектом с последующим извлечением терминов [2] из них. Данные системы подразумевают наличие диалога с субъектом с целью получения развернутых текстовых ответов на поставленные вопросы. Далее из полученных от субъекта текстов извлекаются термины.

Оба аналога имеют соответственно недостатки:

1. Проблематично использование готовых словарей для недавно сформированных или ряда узкоспециальных предметных областей.
2. Не факт, что термины из текстов, подлежащих фильтрации, должны присутствовать в текстах, выданных субъектом.

Для преодоления критики аналогов было предложено накапливать термины, извлекая их из текстов, подлежащих фильтрации. При этом существенным требованием к накопленному списку терминов является его достаточность. В настоящее время в качестве критерия достаточности накопленного списка рассматривают малое (ниже порогового значения) количество вновь прибывших терминов [2]. Данный критерий не учитывает возможную фоносемантическую близость терминов [3], особенно при работе с достаточно узкой предметной областью, и, таким образом, может оказаться избыточным.

С учетом сказанного было предложено формировать нормированные хеш-коды терминов [4], а затем исследовать совокупность кодов накопленных терминов на близость к генеральной [5] в рамках рассматриваемой предметной области.

1. Леонтьев Д.А. Психология смысла: природа, строение и динамика смысловой реальности, Смысл (2003).
2. Всеволодский С.Н., Гаврилов А.В. Архитектура интеллектуальной системы тестирования знаний с анализом ответов на естественном языке. Международная конференция ИСТ-2003 "Информационные системы и технологии" (2003).
3. Журавлев А.П Звук и смысл : Кн. для внеклас. чтения учащихся ст. классов, Просвещение (1991).
4. Хеширование [Электронный ресурс] // Википедия. URL: <https://ru.wikipedia.org/wiki/Хеширование>
5. Генеральная совокупность и выборочный метод [Электронный ресурс] – Режим доступа: <http://www.grandars.ru/student/statistika/generalnaya-sovokupnost.html>